UNIVERSITY OF MARYLAND
COMPUTER SCIENCE CENTER

COLLEGE PARK, MARYLAND
20742

(15) DAAG53-76-C-Ø138, DARPA Order 3206

(1)

(14) TR-762
DAAG-53-76C-0138

(11) May [redacted]79

(6) SHAPE SEGMENTATION USING RELAXATION.

(12) 31

(10) Wallace S. Rutkowski,
Shmuel Peleg
Azriel Rosenfeld

Computer Science Center
University of Maryland
College Park, MD 20742

(9) Technical rept.,

ABSTRACT

Relaxation is applied to the segmentation of closed boundary curves of shapes. The ambiguous segmentation of the boundary is represented by a directed graph structure whose nodes represent segments, where two nodes are joined by an arc if the segments are consecutive along the boundary. A probability vector is associated with each node; each component of this vector provides an estimate of the probability that the corresponding segment is a particular part of the object. Relaxation is used to eliminate impossible sequences of parts, or reduce the probabilities of unlikely ones. In experiments involving airplane shapes, this almost always results in a drastic simplification of the graph, with only good interpretations surviving.

The support of the Defense Advanced Research Projects Agency and the U.S. Army Night Vision Laboratory under Contract DAAG-53-76C-0138 (DARPA Order 3206) is gratefully acknowledged, as is the help of Kathryn Riley in preparing this paper.

403 018

1. Introduction

In many image processing tasks it is important to segment an object into subparts for subsequent processing. For some problems this is easily accomplished by a single, simple level of processing, such as breaking the boundary of the object at corners whose angles are sharper than a given threshold. In other cases, it is necessary to employ some knowledge of the objects' structure to obtain a correct segmentation. For example, attributed grammars are used for boundary analysis in [1]. A general survey of shape analysis can be found in [2].

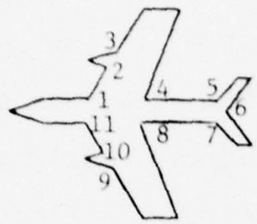
This paper describes an application of relaxation labeling [3,4] to the segmentation of closed object boundary curves. An object boundary is first segmented ambiguously, using very liberal segmentation criteria. This ambiguous segmentation is represented by a directed graph structure which contains (hopefully) the desired segmentation as a subgraph. The nodes of the graph represent pieces of the object boundary. Each node has associated with it a probability vector, assigned by a classification program, representing the segment's probability of being various parts of the object. Node A is connected to node B by a directed arc iff the boundary segment represented by B directly follows the segment represented by A. Since only certain sequences of parts are allowable, we can apply a relaxation process to prune the graph structure. Our experiments show that the relaxation process usually results in a drastic simplification of the graph, which greatly simplifies the problem of extracting desired subgraphs by a sequential process.

It is planned to extend the above methodology in the following ways:

1) Use of an ambiguous segmentation which includes gap filling and shape completion to handle occluded objects.
2) Introduction of relations other than the "follows" relation.
3) Representation by a hierarchical graph structure.

2. Ambiguous segmentation and relaxation

Suppose we have an (x,y) coordinate pair representation of a closed curve, as illustrated below:



Certain distinguished points are called "segmentation points" and are assigned labels (in this case the numbers 1-11). Any piece of contour which begins and ends with a segmentation point will be called a segment. Obviously any segment is uniquely identified by an ordered pair consisting of the labels of the beginning and ending segmentation points, and segments will be thus identified in the remainder of this report.

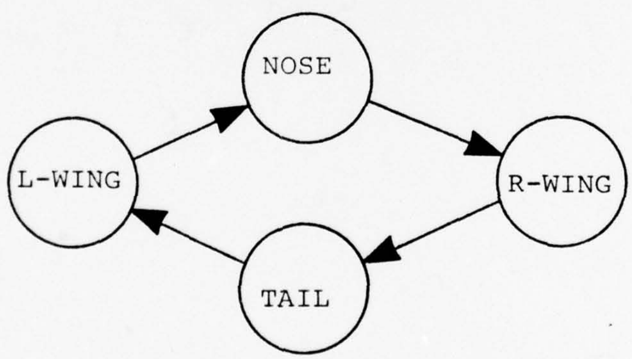
A segmentation of the contour is any set of non-overlapping segments that cover the entire contour. Our objective is, given an object contour, to construct a segmentation such that the segments cover the contour without overlapping, and the segments can be assigned semantically meaningful labels (nose, tail, etc. in the case of an airplane).

In some cases, a simple means of segmentation will suffice. We could, for example, define the segmentation points to be large concave angles, and take as segments any piece of curve lying between two successive segmentation points. In many real-world examples, however, a simplistic strategy will fail.

Our method beings by forming segments using a liberal criterion. Hopefully, the segments thus obtained will contain the desired segmentation (or most of it) as a subset. We then form a directed graph in which the nodes are the segments, and segment $[i,j]$ is connected to segment $[k,\ell]$ by a "follows" relation iff $j=k$ (i.e., $[i,j]$ follows $[k,\ell]$ in the graph iff $[k,\ell]$ follows $[i,j]$ on the boundary). We then classify the segments, assigning to each segment a probability vector, each component of which represents the probability the segment is a certain object part. Specifically, in our airplane example, the probability vector is a four-tuple (P(nose), P(right wing), P(tail), P(left wing)).

Given such a graph structure, with probability vectors associated with the nodes, we can apply a relaxation process to reinforce consistent segments and eliminate inconsistent ones. This method was also used in [4] for ambiguously segmented handwriting. Briefly, the relaxation process operates in the following way. Each node N has a set of predecessor nodes and a set of successor nodes. Consider all triples of labels $\ell_P \ell_N \ell_S$ where ℓ_S is a label associated with a predecessor node, ℓ_N is a label of N, and ℓ_P a label of a successor node. Only certain triples form permissible sequences of labels. We can then modify the probability of each label, depending on the probabilities of the permissible sequences in which it occurs. Thus a label which does not participate in any allowable sequence will have its probability set to zero, while a label which occurs in one or more allowable sequences with high probability will have its probability increased.

After the relaxation process, we search the remaining graph for four-cycles. Each of the four nodes in such a cycle may have one or more possible labels remaining. We can define an interpretation of the cycle to be any assignment of a unique label to each node, where the label assigned to each node is among the remaining labels for that node. A legal interpretation is any interpretation in which the labels in the cycle follow each other in an allowable order, in our case:

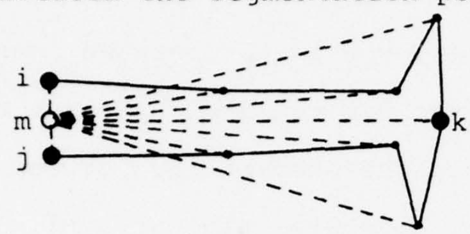


Each legal interpretation is then assigned a figure of merit as follows. With each node in the four-cycle we have associated a unique label. We now look back at the original probability vector for this node to find the original probability of this label. The probabilities thus found at each node are multiplied, and the result is the figure of merit. We can then select as our final result the interpretation with the highest figure of merit.

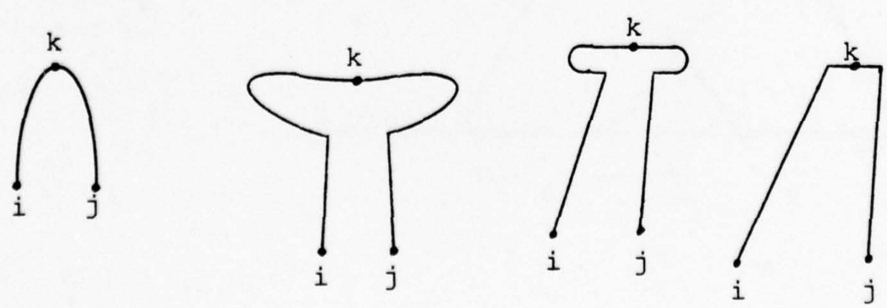
Of course, the use of four-cycles is specific to our problem. More generally, one would conduct a state space search on the resulting graph to find legal interpretations which are not necessarily of fixed length.

3. Assignment of the initial probabilities

Suppose we have a segment [i,j] to which we wish to assign a probability vector (N,R,L,T). Our program first finds a point, k, between the segmentation points i and j as follows:



Consider a polygonal approximation of the curve segment [i,j] where the vertices of the polygon are placed at the x,y coordinates associated with the local maxima and minima of the curvature. We then define point m to be midway between i and j. We can now examine the line segments radiating from m to the vertices of the polygon. We discard any such line segment which does not lie entirely within the shape. (Specifically, if we are considering line segment $m\ell$, all vertices which precede ℓ must lie to the left of line segment $m\ell$, and all vertices which succeed ℓ must lie to the right of $m\ell$.) We then select as k that point which miximizes the length of the line segment. The location of k for several wing, nose, and tail shapes is illustrated below:



Now, using k, the following operations were performed:

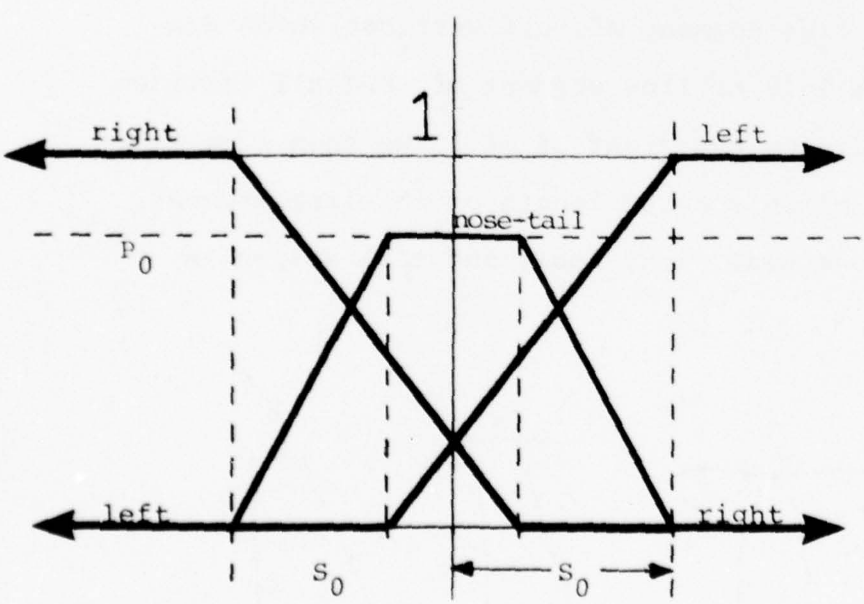
1) The angle between line segments ik and kj was measured. If this angle was too shallow the segment [i,j] was discarded.

2) A crude measure of symmetry was computed as:

sym = (perimeter between k and j) - (perimeter between i and k).

3) The probability vector was computed by first using the value of sym to distribute the probability between three probabilities, nose-tail (low absolute value of sym), left wing (high positive value of sym), and right wing (high negative value of sym). Then, if k was a local maximum, the probability of nose-tail was assigned to nose; if k was a local minimum, the probability of nose-tail was assigned to tail.

The probabilities were distributed according to the scheme below:



P_0 = Prob(a perfectly symmetrical object is a nose or tail)

S_0 = maximum value of symmetry at which an object can have a non-zero probability of being a nose or tail.

4. Experimental results

In Figure 1 we see an airplane contour with a plot of its curvature. The numbers on the figure and on the curvature plot indicate some corresponding points. The curvature measure used was the method of weighted k-curvature discussed in [5]. Table 1.1 shows the results of the segmentation and classification for this figure. Table 1.2 displays the effect of relaxation. Of particular significance is the sharp reduction in the number of 4-cycles and the number of interpretations. It is these numbers which provide a measure of the size of the space which must be serached to find the legal interpretations. Table 1.3 shows the surviving nodes and labels after three iterations of relaxation, and finally, Table 1.4 shows the legal interpretations and their merit. These same interpretations could have been found in the graph before the relaxation process; however, before relaxation we would have to find the 3 legal interpretations among the 222 possible interpretations, whereas after only 3 iterations of relaxation we have reduced the search space to 10 possible interpretations.

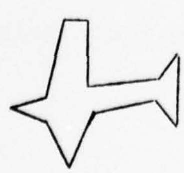
The merit value is obtained by multiplying together the original probability estimates of the labels, and scaling the resulting numbers upward to facilitate comparison. In Table 1.4, we find that cycle 2 has the highest merit, and if we check the interpretation with the contour, we see that this is the correct segmentation. (For example, in cycle 2, the right wing is node 7, which we see from Table 1 or 3 has the range 4-12. In the airplane figure, the right wing does indeed go from point 4 to point 12.) The fact that cycle 3 has such a relatively high merit is due in part to the shortcoming of our classification program, which allowed some probability that segment 9 might be a nose, and in part to the fact that by using only the "follows" relation, our process is totally insensitive to the gross lack of symmetry of cycle 3.

Two other examples are given in Figures 2-3 and Tables 2.1-4 and 3.1-4.

5. Future research

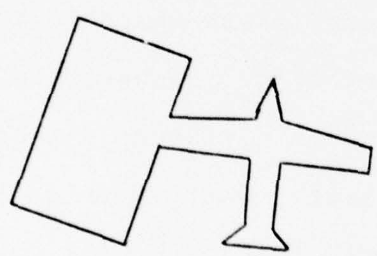
One obvious criticism of the method as currently implemented is the use of problem-specific features and the ad hoc manner in which they are found. This is a common problem in classical pattern recognition. Part of the problem in our case is that in this preliminary study the airplane shape was broken into only four sub-parts, some of which can have a significant amount of structural detail. We could break the airplane shape into more sub-parts so that each resulting part would have a simpler structure, and could be classified initially by some more general technique. There is also another, more fundamental problem here. It would be desirable, of course, to have a universal feature set, in terms of which an arbitrary shape could be described and recognized. At the present time, however, the nature of such a universal feature set is a significant research area in need of much further investigation. Until more concrete ideas about such a feature set are forthcoming, we will continue to face the problem of features which are specific to the problem domain and somewhat ad hoc in nature.

Another clear limitation of the technique is that presently the only relationship used is the "follows" relation. The program would, for example, accept the shape below as an airplane:

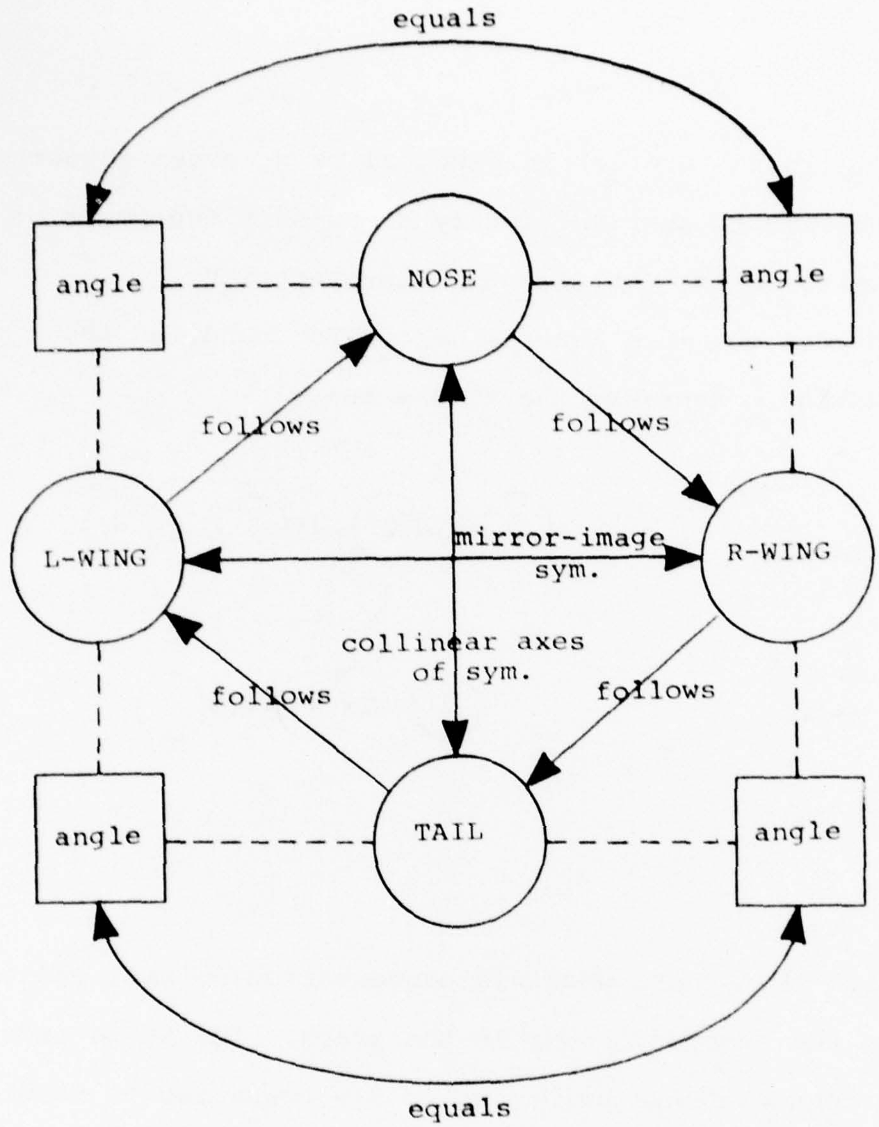


We could reject this shape by using other relations, such as requiring that the two wings be similar in shape, and that the axes of symmetry of the nose and tail sections be collinear. Of course, in doing this, we will be representing the airplane by a more general graph structure than the simple closed cycle, and finding the plane will become a more general problem of subgraph finding. There would seem to be a good application here for Kitchen's work [6] on the use of relaxation in subgraph matching.

Another situation not handled by the current program is the problem of missing or occluded pieces of shape. We can give the program some limited capacity in this area. Consider the shape below:

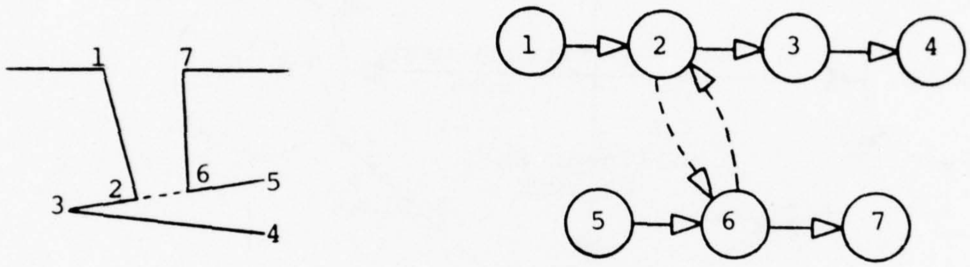


Although the program will not find the left wing, it will be able to find a nose, followed by a right wing, followed by a tail. Thus the right wing will certainly survive the first iteration of relaxation, and, if we allow the nose and tail nodes to survive, even though supported on only one side, then the nose-right wing-tail subgraph could survive. However, some preliminary experiments with this technique indicate that when


equals
angle
NOSE
angle
follows
follows
L-WING
mirror-image
sym.
R-WING
collinear axes
of sym.
follows
follows
angle
TAIL
angle
equals

only the "follows" relation is used, allowing nodes supported on only one side to survive greatly diminishes the power of the program to filter out illegal interpretations.

We can create a more general method for handling the above situation. Consider the figure below:



The numbered points are possible segmentation points, and correspond to the numbered nodes in the graph. The solid arcs in the graph connect those points which are connected to each other by visible portions of object contour. The dashed arcs connect points connected by proposed gap completions [7]. If we consider a graph with only the solid arcs, we can uniquely specify any segment of an ambiguous segmentation by giving the starting and ending node. In the more general case, where gap completions are allowed, the segments of an ambiguous segmentation would be specified as sequences of nodes representing possible paths through the graph (i.e., (1,2,6,7) or (5,6,2,3,4)).

Suppose that we use other relations in addition to the "follows" relation. Specifically, suppose that we require the two wings to be similar, and the axes of symmetry of the nose and tail to be collinear, as suggested earlier. Then the shape below might still be accepted as an airplane:

The asymmetry introduced into the tail by the different angles of the wings is very slight, and we would not want to reject the shape based on this alone. What we can do is measure the angle between each wing and the nose (and tail), and demand that these angles be similar (see the figure on the next page). These angles can be regarded as second level features in a hierarchical graph, since these nodes are created by combining information from several (in this case two) nodes at the first level of the graph.

The use of a hierarchical structure offers a number of advantages:

1) Certain aspects of the problem (and many other computer vision problems) lend themselves naturally to a hierarchical description. It is common to describe a scene in terms of parts,

which may themselves have subparts, etc. A hierarchical structure also provides a convenient way of thinking about the use of higher level features that stem from the relationships of image parts to each other, such as relative angle and distance. We can use these features to express relationships such as "A is farther from B than B is from C."

2) A hierarchical structure can speed the rate of information propagation. In a single-layer relaxation network, information propagates in linear time, i.e., the time taken for a node to receive information from a node n links distant is proportional to n. With a tree or pyramid structure, information can propagate in log n time.

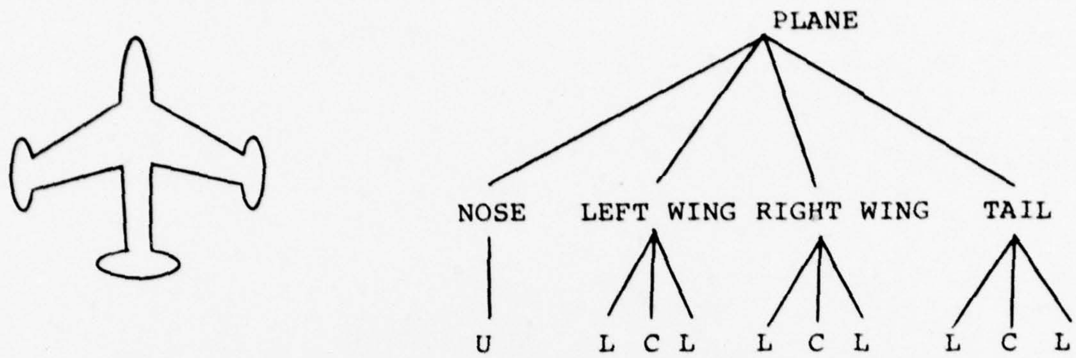
3) A hierarchy can reduce the number of labels required. We may find, for example, that many of the parts into which we wish to segment an object have similar shapes. It would be preferable to have a single label to apply to all of these parts, rather than a separate label for each occurrence of the part in the object. For example, given the three shapes below,

L (line segment)

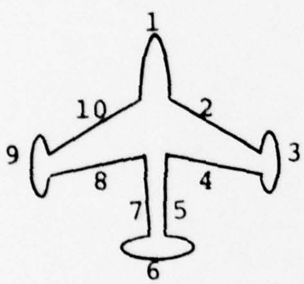
U ("U" shape)

C ("C" shape)

we could represent an airplane model as:



in a hierarchy, rather than with a unique label for every part at the lowest level:



1) nose
2) right leading edge
3) right tip tank
4) right trailing edge
5) right fuselage
6) tail
7) left fuselage
8) left trailing edge
9) left tip tank
10) left leading edge

Some previous work on the application of a hierarchical structure to ambiguous segmentation was done by L. Davis [8] for the case of one-dimensional waveforms.

6. Concluding remarks

Our experiments have verified the ability of relaxation to significantly reduce the complexity of the segmentation graph. The number of interpretations which must be searched is typically reduced by one or two orders of magnitude. The combination of ambiguous segmentation and a graph structure representation of the shape allows our program to correctly segment a wide class of ambiguous shapes, in a manner not possible with the use of template matching techniques. The program thus demonstrates the important characteristic of using a more general, abstract representation of an airplane shape than that defined by a prototype. The basic methodology used can be extended to more complex graph structures utilizing shape completion, a variety of object part relations, and a hierarchical structure. These extensions provide promising avenues for future research.

References

1. K. C. You and K. S. Fu, Syntactic shape recognition using attributed grammars, in R. A. Kirsch and R. N. Nagel, eds., Eighth Annual Automatic Imagery Pattern Recognition Symposium Proceedings (Gaithersburg, MD, April 1978).

2. T. Pavlidis, A review of algorithms for shape analysis, Computer Graphics and Image Processing 7, 1978, 243-258.

3. S. Peleg, A new probabilistic relaxation scheme, TR-711, Computer Science Technical Report, University of Maryland, November 1978; to appear in 1979 Pattern Recognition and Image Processing Conference, Chicago, IL, August 1979.

4. S. Peleg, Ambiguity reduction in handwriting with ambiguous segmentation and uncertain interpretation, Computer Graphics and Image Processing, in press.

5. W. S. Rutkowski and A. Rosenfeld, A comparison of corner-detection techniques for chain-coded curves, TR-623, Computer Science Technical Report, University of Maryland, January 1978.

6. L. Kitchen, Discrete relaxation for matching relational structures, TR-665, Computer Science Technical Report, University of Maryland, June 1978.

7. W. S. Rutkowski, Shape completion, Computer Graphics and Image Processing 9, 1979, 89-101.

8. L. S. Davis and A. Rosenfeld, Hierarchical relaxation for waveform parsing, in A. Hanson and E. M. Riseman, eds., Computer Vision Systems, Academic Press, 1978, 101-109.

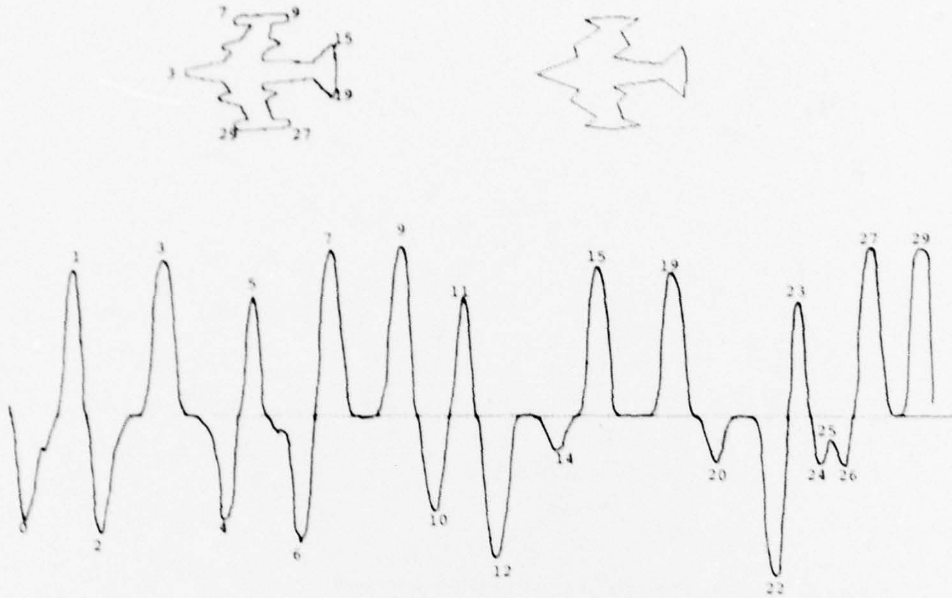


Figure 1. An airplane contour and its curvature plot. Some of the peaks are numbered to show correpondence with the image and the segmentation tables.

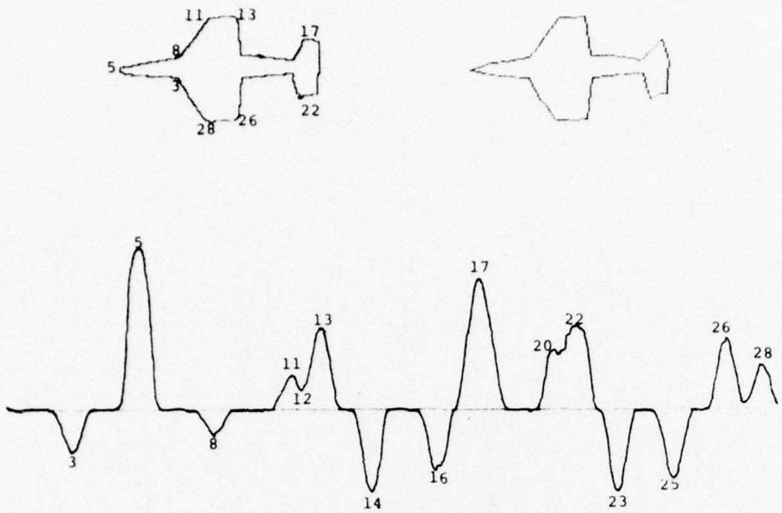


Figure 2. Same as Figure 1, but for another airplane.

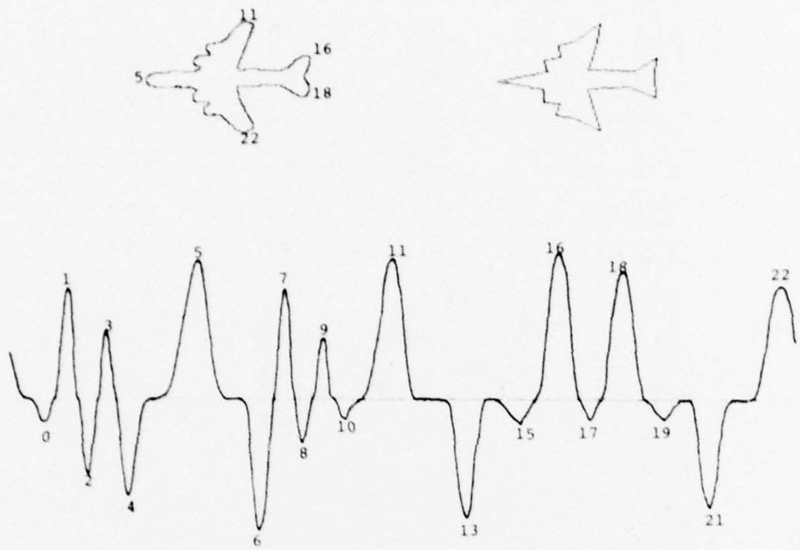


Figure 3. Same as Figure 1, but for another airplane.

SEGMENT NUMBER	RANGE OF SEGMENT	NOSE	PROBABILITIES RWING	TAIL	LWING
1	0-4	--	1	--	--
2	0-6	.6	.22	--	.18
3	2-4	.6	.22	--	.18
4	2-6	--	--	--	1
5	2-10	--	--	--	1
6	4-10	--	1	--	--
7	4-12	--	.32	.6	.08
8	4-14	--	--	--	1
9	6-10	.04	.96	--	--
10	6-12	.6	.26	--	.14
11	10-22	--	.98	.02	--
12	12-20	--	1	--	--
13	12-22	--	--	.52	.48
14	12-24	--	--	--	1
15	12-25	--	--	--	1
16	14-20	--	--	--	1
17	14-22	--	--	--	1
18	22-0	--	1	--	--
19	22-2	--	.08	.6	.32
20	24-0	--	1	--	--
21	24-2	--	--	.16	.84
22	24-4	--	1	--	--
23	25-0	--	1	--	--
24	25-2	--	--	.06	.94
25	25-4	--	1	--	--
26	26-0	--	--	--	1
27	26-2	--	--	--	1
28	26-4	--	1	--	--

Table 1.1. Ambiguous segmentation of Figure 1. Segments are between two peaks.

ITERATION NUMBER	NUMBER OF NODES	NUMBER OF LABELS	NUMBER OF 4-CYCLES	NUMBER OF INTERPRETATIONS
0	28	43	118	222
1	10	15	12	20
2	9	12	8	16
3	8	10	6	10

Table 1.2. Reduction of ambiguity for the airplane of Figure 1 in 3 iterations of relaxation.

SEGMENT NUMBER	RANGE	LABEL
3	2-4	NOSE
4	2-6	LWING
6	4-10	RWING
7	4-12	RWING
9	6-10	NOSE
11	10-22	RWING = .26, TAIL = .74
13	12-22	TAIL
19	22-2	LWING = .5, TAIL = .5

Table 1.3. The ambiguous segmentation of Figure 1 after 3 iterations of relaxation. This table is analogous to Table 1.1.

CYCLE NUMBER	NOSE NODE	RWING NODE	TAIL NODE	LWING NODE	MERIT
1	3	6	11	19	3.8
2	3	7	13	19	31.9
3	9	11	19	4	23.5

Table 1.4. The cycles surviving from Figure 1 after 3 iterations of relaxation. The nodes are segments from Table 1.1, and the merit was computed by using the original probabilities.

SEGMENT NUMBER	RANGE OF SEGMENT	PROBABILITIES NOSE	RWING	TAIL	LWING
1	3-8	.6	.12	--	.28
2	3-14	--	--	--	1
3	8-14	.2	.8	--	--
4	14-23	--	1	--	--
5	14-25	--	--	.6	.4
6	16-23	.16	.84	--	--
7	16-25	--	--	--	1
8	23-3	.02	.98	--	--
9	25-3	.1	--	--	.9
10	25-8	--	1	--	--

Table 2.1 Ambiguous segmentation of Figure 2.

ITERATION NUMBER	NUMBER OF NODES	NUMBER OF LABELS	NUMBER OF 4-CYCLES	NUMBER OF INTERPRETATIONS
0	10	17	7	36
1	4	6	1	4
2	4	4	1	1

Table 2.2. Relaxation results.

SEGMENT NUMBER	RANGE	LABEL
1	3-8	NOSE
3	8-14	RWING
5	14-25	TAIL
9	25-3	LWING

Table 2.3. Nodes surviving third iteration.

CYCLE NUMBER	NOSE NODE	RWING NODE	TAIL NODE	LWING NODE	MERIT
1	1	3	5	9	259.2

Table 2.4. Allowable interpretations with their merits.

SEGMENT NUMBER	RANGE OF SEGMENT	NOSE	PROBABILITIES RWING	TAIL	LWING
1	0-8	.28	.72	--	--
2	0-10	.6	.24	--	.16
3	2-6	.3	.7	--	--
4	2-8	.6	.2	--	.2
5	2-10	.32	--	--	.68
6	4-6	.6	.24	--	.16
7	4-8	.34	--	--	.66
8	4-13	--	1	--	--
9	6-13	.16	.84	--	--
10	6-15	--	--	--	1
11	6-17	--	--	--	1
12	8-13	.6	.32	--	.06
13	10-13	.46	--	--	.54
14	13-19	--	1	--	--
15	13-21	--	.22	.6	.18
16	15-19	--	--	--	1
17	15-21	--	--	--	1
18	17-19	.6	.14	--	.26
19	17-4	--	1	--	--
20	19-4	.56	--	--	.44
21	21-0	.58	.42	--	--
22	21-2	.5	--	--	.5
23	21-4	.04	--	--	.96
24	21-6	--	1	--	--

Table 3.1 Ambiguous segmentation of Figure 3.

ITERATION NUMBER	NUMBER OF NODES	NUMBER OF LABELS	NUMBER OF 4-CYCLES	NUMBER OF INTERPRETATIONS
0	24	46	153	304
1	13	17	20	16
2	10	10	20	4
3	9	9	16	4

Table 3.2 Relaxation results.

SEGMENT NUMBER	RANGE	LABEL
3	2-6	NOSE
4	2-8	NOSE
6	4-6	NOSE
7	4-8	NOSE
9	6-13	RWING
12	8-13	RWING
15	13-21	TAIL
22	21-2	LWING
23	21-4	LWING

Table 3.3. Nodes surviving third iteration.

CYCLE NUMBER	NOSE NODE	RWING NODE	TAIL NODE	LWING NODE	MERIT
1	3	9	15	22	75.6
2	4	12	15	22	61.2
3	6	9	15	23	290.3
4	7	12	15	23	66.6

Table 3.4. Allowable interpretations with their merits.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
SHAPE SEGMENTATION USING RELAXATION

5. TYPE OF REPORT & PERIOD COVERED
Technical

6. PERFORMING ORG. REPORT NUMBER
TR-762

7. AUTHOR(s)
Wallace S. Rutkowski
Shmuel Peleg
Azriel Rosenfeld

8. CONTRACT OR GRANT NUMBER(s)
DAAG-53-76C-0138

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Computer Science Center
University of Maryland
College Park, MD 20742

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
U.S. Army Night Vision Laboratory
Fort Belvoir, VA 22060

12. REPORT DATE
May 1979

13. NUMBER OF PAGES
27

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Pattern recognition
Shape analysis
Segmentation
Relaxation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
Relaxation is applied to the segmentation of closed boundary curves of shapes. The ambiguous segmentation of the boundary is represented by a directed graph structure whose nodes represent segments, where two nodes are joined by an arc if the segments are consecutive along the boundary. A probability vector is associated with each node; each component of this vector provides an estimate of the probability that the corresponding segment is a particular part of the object. Relaxation is used to eliminate impossible

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) sequences of parts, or reduce the probabilities of unlikely ones. In experiments involving airplane shapes, this almost always results in a drastic simplification of the graph, with only good interpretations surviving.